

SADDLEPOINT APPROXIMATIONS TO THE DISTRIBUTION OF THE TOTAL DISTANCE OF THE MULTIVARIATE ISOTROPIC AND VON MISES-FISHER RANDOM WALKS

Riccardo Gatto

Submitted: June 2016

Revised: October 2016

Abstract

This article considers the random walk over \mathbb{R}^p , with $p \geq 2$, where the directions taken by the individual steps follow either the isotropic or the von Mises-Fisher distributions. Saddlepoint approximations to the density and to upper tail probabilities of the total distance covered by the random walk, i.e. of the length of the resultant, are derived. The saddlepoint approximations are one-dimensional and simple to compute, even though the initial problem is p -dimensional. Numerical illustrations of the high accuracy of the proposed approximations are provided.

Key words and phrases

Bessel function; directional distribution; large deviations approximation; relative error; resultant length.

The author thanks an anonymous Referee, the Deputy Editor and the Editor-in-Chief for helpful comments and suggestions.

2010 Mathematics Subject Classification

41A60 Asymptotic approximations, asymptotic expansions (steepest descent, etc.)

62H11 Directional data; spatial statistics

Address

Institute of Mathematical Statistics and Actuarial Science

Department of Mathematics and Statistics

University of Bern

Alpeneggstrasse 22, 3012 Bern, Switzerland

+41 31 6318807

gatto@stat.unibe.ch www.stat.unibe.ch/gatto

orcid.org/0000-0001-8374-6964

1 Introduction

This article considers the random walk over the Euclidean space \mathbb{R}^p , for unrestricted $p = 2, 3, \dots$, where the directions taken by the individual steps are generated from a prescribed directional distribution. A directional distribution is a probability distribution over the unit sphere and a directional random vector has length one. The corresponding terms for the case $p = 2$ are circular distribution and circular random vector. A directional distribution possesses no absolute zero and so any designation of large or small values becomes purely arbitrary. The major reference is perhaps Section 9.3 of Mardia and Jupp (2000).

The precise contribution of this article is the derivation of saddlepoint approximations to the density and to upper tail probabilities of the total distance run through by the random walk, with isotropic or von Mises-Fisher directions and with arbitrary dimension p . Numerical illustrations of the accuracy of the saddlepoint approximations are also provided. Some related articles are the following. Barakat (1973) proposed a computational scheme for a Fourier series approximation to the density of the total distance of the circular isotropic random walk, with uniformly distributed step size. Exploiting results on Bessel functions, Orsingher and De Gregorio (2007) obtained the distribution of the random walk with isotropic directions and exponentially distributed step lengths. Alternative results can be found in Stadje (1989) and Masoliver et al. (1993). Saddlepoint approximations for the isotropic and von Mises planar random walks, are given in Jensen (1995), p. 162-165, and Gatto (2017), who considered constant and exponential step lengths, as well as inhomogeneous birth processes for the number of steps. Further, Weiss and Kiefer (1983) proposed a saddlepoint approximation for the projection of the circular isotropic random walk with fixed but unequal step sizes. Gatto and Mayer (2005) proposed a saddlepoint approximation for this projected random walk, but with exponential step size. Moreover, Gatto and Jammalamadaka (2003) proposed a saddlepoint approximation for the circular random walk with wrapped symmetric α -stable directions.

The saddlepoint expansion to the density of the mean of n independent and identically distributed (i.i.d.) random variables was suggested by the fundamental article of Daniels (1954). It is an asymptotic expansion in powers of n^{-1} , as $n \rightarrow \infty$, with the saddlepoint approximation as leading term. The saddlepoint approximation possesses a relative error of the order n^{-1} , at any point of the support of the density, that is over the large deviations region of the mean. In comparison, the normal approximation possesses an absolute error of the order $n^{-1/2}$ only and its validity is restricted to the normal deviations region. The saddlepoint approximation to the upper tail probability was suggested by Lugannani and Rice (1980). The saddlepoint approximation is thus a large deviations technique, substantially more accurate normal or Edgeworth approximations, especially when computing very small tail probabilities and even with very small samples sizes. Saddlepoint approximations compete with techniques of rare event simulation, importance sampling essentially, but they do not require Monte Carlo sampling.

This random walk has been considered in various physical sciences, see e.g. Barber and Ninham (1970), including statistical mechanics, see e.g. Flory (1969), crystallography, see

e.g. Srinivisan and Parthasarathy (1976) and Chen et al. (2015), atmospheric science, etc. In statistics, the total distance covered by the random walk is relevant to inference on the underlying directional concentration. So the saddlepoint approximations derived in this article can be used for computing P-values, critical regions and power functions of the likelihood ratio test of isotropy against alternative von Mises Fisher distributions; more details are given in the last paragraph of Section 2.

Saddlepoint approximations represent a classical theme of mathematical statistics and a short review is the following. Field (1982) suggested a saddlepoint approximation for multivariate M-estimators. Feuerverger (1989), Wang (1992) and Ronchetti and Welsh (1994) proposed the saddlepoint approximation in the context of bootstrap inference, i.e. when the underlying distribution of the is replaced by the empirical distribution of the sample. Wang (1995) proposed an iteration of the saddlepoint approximation yielding an approximation for quantiles, essential for computing critical regions of tests. Wang (1993) considered saddlepoint approximations for finite populations problems, arising in survey sampling. Daniels and Young (1991) provided a saddlepoint approximation for the studentized mean. Gatto and Jammalamadaka (1999) proposed a saddlepoint approximation for important classes of non-parametric tests, including one-sample tests based on spacings and two-sample tests based on spacing-frequencies. These two types of tests play a central role in circular data, because they neither depend on the null direction nor on the (clockwise or counter-clockwise) sense of rotation. Robinson (1982) gave a saddlepoint approximation for one- and two-sample permutation tests. Robinson, Ronchetti and Young (2003) considered multiparameter tests and provided a particular test statistic, based on the exponent of the saddlepoint approximation to the density of the multivariate M-estimator, which has a chi-squared distribution up to a bounded relative error. An application of this saddlepoint test statistic in functional measurement error models is given in Ma and Ronchetti (2011). Jeganathan, Paige and Trindade (2015) used the saddlepoint approximation for bootstrap inference on the spatial dependence parameter, in Gaussian lattice regression models. Some contributions restricted to statistical inference with directional data are the following. Gatto (2000) obtained P-values for the test on the parameters of the wrapped normal model, using the multivariate test of Robinson, Ronchetti and Young (2003). Kume and Wood (2005) suggested saddlepoint approximations to the normalizing constants of the Bingham and Fisher-Bingham distributions. Gatto and Mayer (2005) proposed saddlepoint approximations for some optimal tests regarding the concentration parameter of wrapped symmetric α -stable models. A general reference on the saddlepoint approximation is de Bruijn (1982), Chapter 5, and further main references with orientation in probability and statistics are Section 6.5 of Barndorff-Nielsen and Cox (1989), Field and Ronchetti (1990), Jensen (1995) and Kolassa (2006). A review article for saddlepoint approximations in statistics is Reid (1988) and a review for applications in robust statistics is Field and Tingley (1997).

The remaining part of this article has the following structure. Section 2 recalls the essential facts of spherical measures. Section 3.2 derives the saddlepoint approximations to the density and to upper tail probabilities of the resultant length of the isotropic random walk. Section 3.3 extends the approximations to the random walk under von Mises-Fisher direc-

tions. Finally, Section 4 illustrates the numerical accuracy of the proposed approximations.

2 Basic notions and notation

This section mainly recalls some basic facts regarding spherical measures and introduces the notation. One can re-express $\underline{x} \in \mathbb{R}^p \setminus \{0\}$ in terms of angular coordinates by the diffeomorphism $\underline{x} = \underline{g}(\alpha_1, \dots, \alpha_{p-1}, r)$ which, for $p \geq 3$, is given by

$$\begin{aligned} x_1 &= r \sin \alpha_1 \sin \alpha_2 \dots \sin \alpha_{p-2} \cos \alpha_{p-1}, \\ x_2 &= r \sin \alpha_1 \sin \alpha_2 \dots \sin \alpha_{p-2} \sin \alpha_{p-1}, \\ &\vdots \\ x_{p-1} &= r \sin \alpha_1 \sin \alpha_2, \\ x_p &= r \cos \alpha_1, \end{aligned} \tag{1}$$

where $r > 0$, $0 \leq \alpha_j \leq \pi$, for $j = 1, \dots, p-2$, and $0 \leq \alpha_{p-1} < 2\pi$. For $p \geq 3$, the determinant of the Jacobian matrix of \underline{g} is given by

$$J(\alpha_1, \dots, \alpha_{p-2}, r) = r^{p-1} \prod_{j=2}^{p-1} \sin^{p-j} \alpha_{j-1}. \tag{2}$$

For $p = 2$, this determinant is equal to r . Let $r > 0$ and define the sphere $\mathbb{S}_r^{p-1} = \{\underline{x} \in \mathbb{R}^p \mid \|\underline{x}\| = r\}$. For any continuous (and thus Borel) function $f : \mathbb{S}_r^{p-1} \rightarrow \mathbb{R}$,

$$\int_{\mathbb{S}_r^{p-1}} f(\underline{x}) d\lambda_{p,r}(\underline{x}) = \int_0^{2\pi} \int_0^\pi \dots \int_0^\pi f(\underline{g}(\alpha_1, \dots, \alpha_{p-1}, r)) J(\alpha_1, \dots, \alpha_{p-2}, r) d\alpha_1 \dots d\alpha_{p-1}, \tag{3}$$

where $d\lambda_{p,r}(\underline{x})$ denotes the infinitesimal surface area, or Lebesgue measure, around $\underline{x} \in \mathbb{S}_r^{p-1}$ and over \mathbb{S}_r^{p-1} . Thus the surface area of \mathbb{S}_r^{p-1} is given by

$$a_{p,r} = \int_{\mathbb{S}_r^{p-1}} d\lambda_{p,r}(\underline{x}) = r^{p-1} \frac{2\pi^{\frac{p}{2}}}{\Gamma(\frac{p}{2})}. \tag{4}$$

The random walks considered in this article are directed either by the isotropic distribution or by the von Mises-Fisher distribution over $(\mathbb{S}^{p-1}, \mathcal{B}(\mathbb{S}^{p-1}))$, where $\mathbb{S}^{p-1} = \mathbb{S}_1^{p-1}$ and $\mathcal{B}(\mathbb{S}^{p-1}) = \{B \cap \mathbb{S}^{p-1} \mid B \in \mathcal{B}(\mathbb{R}^p)\}$. The isotropic distribution or uniform distribution is given by

$$U_p[B] = \int_B a_p^{-1} d\lambda_p(\underline{x}) = \frac{\Gamma(\frac{p}{2})}{2\pi^{\frac{p}{2}}} \lambda_p[B], \quad \forall B \in \mathcal{B}(\mathbb{S}^{p-1}),$$

where $a_p = a_{p,1}$ and $\lambda_p = \lambda_{p,1}$. For example, the isotropic density over \mathbb{S}^2 with respect to (w.r.t.) λ_2 is $(4\pi)^{-1}$, because $dU_2(\underline{x}) = \Gamma(3/2)/(2\pi^{3/2}) d\lambda_2(\underline{x}) = (4\pi)^{-1} d\lambda_2(\underline{x})$, $\forall \underline{x} \in \mathbb{S}^1$. The isotropic distribution is primeval in directional data analysis: it represents the case where no mean direction or no preferred direction exists. It also represents the state of

maximal uncertainty or entropy. A directional random vector \underline{X} over \mathbb{S}^{p-1} with the isotropic distribution can also be defined as the one for which, for any orthogonal matrix $A \in \mathbb{R}^{p \times p}$ (viz. $A^\top A = AA^\top = I$), $A\underline{X} \sim \underline{X}$ holds.

The von Mises-Fisher distribution with mean direction $\underline{\mu} \in \mathbb{S}^{p-1}$ and concentration $\kappa \geq 0$ has density w.r.t. U_p and at $\underline{x} \in \mathbb{S}^{p-1}$ given by the exponential model

$$f(\underline{x}|\underline{\mu}, \kappa) = c_p(\kappa) \exp\{\kappa \langle \underline{\mu}, \underline{x} \rangle\}, \quad (5)$$

where

$$c_p(\kappa) = \begin{cases} \left(\frac{\kappa}{2}\right)^{\frac{p}{2}-1} \left\{ \Gamma\left(\frac{p}{2}\right) I_{\frac{p}{2}-1}(\kappa) \right\}^{-1}, & \text{if } \kappa > 0, \\ 1, & \text{if } \kappa = 0. \end{cases} \quad (6)$$

As usually, $\forall z \in \mathbb{C}$,

$$I_\nu(z) = \frac{\left(\frac{z}{2}\right)^\nu}{\pi^{\frac{1}{2}} \Gamma(\nu + \frac{1}{2})} \int_0^\pi e^{z \cos \theta} \sin^{2\nu} \theta d\theta = \frac{\left(\frac{z}{2}\right)^\nu}{\pi^{\frac{1}{2}} \Gamma(\nu + \frac{1}{2})} \int_{-1}^1 (1-x^2)^{\nu-\frac{1}{2}} e^{zx} dx \quad (7)$$

is the modified Bessel function I of order ν , such that $\Re \nu > -1/2$; see e.g. Abramowitz and Stegun (1965), p. 376. In the circular case, $c_2(\kappa) = I_0^{-1}(\kappa)$ and the density expressed in terms of the angle w.r.t. a null direction is the von Mises density. Note that the density w.r.t. to the Lebesgue measure λ_p would have normalizing constant $c_p(\kappa)$ replaced by $c_p(\kappa)/a_p$, in (5). We use the abbreviation $\text{vMF}(\underline{\mu}, \kappa)$ distribution for the von Mises-Fisher distribution with mean direction $\underline{\mu}$ and concentration κ . The von Mises-Fisher distribution is important because it possesses many theoretical properties or characterizations, often valid for the normal distribution once transposed from the sphere to the Euclidean space. For instance, the von Mises-Fisher (cf. normal) is a maximum entropy distribution and its directional marginals are again von Mises-Fisher (cf. normal) distributions.

An important function in this context is

$$A_p(u) = \frac{I_{\frac{p}{2}}(u)}{I_{\frac{p}{2}-1}(u)}, \quad \forall u \geq 0.$$

Note that A_p is a continuous increasing mapping from $[0, \infty)$ onto $[0, 1)$. First, it follows directly from (7) that $A_p(0) = 0$. Second, from

$$I_\nu(u) = \frac{e^u}{\sqrt{2\pi u}} \left\{ 1 - \frac{4\nu^2 - 1}{8u} + O(u^{-2}) \right\}, \quad \text{as } u \rightarrow \infty, \quad (8)$$

cf. Abramowitz and Stegun (1965), p. 377, one deduces that $\lim_{u \rightarrow \infty} A_p(u) = 1$. Third, Amos (1974) shows that $A'_p(u) > 0$, $\forall u > 0$, and so A_p is increasing. Thus A_p is a continuous probability distribution function over $[0, \infty)$.

Let $n \in \mathbb{N} \setminus \{0\}$ and define $\underline{X}_1, \dots, \underline{X}_n$ as independent and $\text{vMF}(\underline{\mu}, \kappa)$ distributed. The resultant vector is denoted $\underline{R}_n = \sum_{j=1}^n \underline{X}_j$, its length $R_n = \|\underline{R}_n\|$ and its direction $\underline{M}_n = \underline{R}_n/R_n$. It can be seen that the maximum likelihood estimators of $\underline{\mu}$ and κ are given by \underline{M}_n and $A_p^{(-1)}(R_n/n)$, respectively, where $A_p^{(-1)}$ denotes the inverse function i.e. the quantile

function of A_p . Consider the testing problem $H_0: \kappa = 0$ against $H_1: \kappa > 0$, where $\underline{\mu}$ is unknown. The likelihood ratio statistic for this problem is given by

$$\Lambda_n = -2 \log \prod_{j=1}^n \frac{f(\underline{X}_j | \underline{\mu}, 0)}{f(\underline{X}_j | \underline{M}_n, A_p^{(-1)}(\frac{R_n}{n}))} = -2n \log c_p \circ A_p^{(-1)}\left(\frac{R_n}{n}\right) + \frac{2}{n} A_p^{(-1)}\left(\frac{R_n}{n}\right) R_n^2.$$

From (31) follows that Λ_n is an increasing function of R_n and so the test rejects H_0 for large values of R_n . This is indeed Rayleigh's test of isotropy, see e.g. Section 15.6 of Mardia et al. (1979). Thus, the saddlepoint approximations to the distributions of R_n under H_0 and under H_1 provide approximations to P-values and to the power function of this test.

3 Total distance of the random walk under isotropic or von Mises-Fisher directions

The random walk with isotropic step directions is considered in Section 3.1 and with von Mises-Fisher step directions in Section 3.2. In these two sections, saddlepoint approximations to the densities and to the upper tail probabilities of the total length of the random walk are derived.

3.1 Isotropic random walk

Let $\underline{X}_1, \dots, \underline{X}_n$ be independent and isotropic over the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $A \in \mathbb{R}^{p \times p}$ be an orthogonal matrix. The notation $V_1 \sim V_2$ means that V_1 and V_2 have same distribution and $V_1 | W_1 \sim V_2 | W_2$ means that the conditional distribution of V_1 given W_1 is equal to the one of V_2 given W_2 . Then $(\underline{X}_1, \dots, \underline{X}_n) \sim (A\underline{X}_1, \dots, A\underline{X}_n)$ implies $(\underline{M}_n, R_n) \sim (A\underline{M}_n, R_n)$, which implies isotropy of \underline{M}_n and also $\underline{M}_n | R_n \sim A\underline{M}_n | R_n$. Thus $\underline{M}_n \sim \underline{M}_n | R_n$ and so \underline{M}_n and R_n are independent. Because the converse holds as well, see Kent et al. (1979), the following lemma holds. Denote by l_p the Lebesgue measure over \mathbb{R}^p .

Lemma 3.1 (Factorization lemma). *Let $n \geq 2$ and $\underline{X}_1, \dots, \underline{X}_n$ be i.i.d. with any absolutely continuous distribution w.r.t. U_p over \mathbb{S}^{p-1} . Let $w_{p,n}$ be the joint density of (\underline{M}_n, R_n) w.r.t. $U_p \times l_1$. Let $q_{p,n}$ be the density of R_n , w.r.t. l_1 . Then \underline{M}_n and R_n are independent, viz.*

$$w_{p,n}(\underline{x}, r) = q_{p,n}(r), \quad \forall \underline{x} \in \mathbb{S}^{p-1}, r \in (0, n],$$

iff the sample values are isotropic.

Consider the polar coordinates of \underline{M}_n as given in (1). Let $w_{p,n}^\circ$ be the joint Lebesgue density of the polar angles $(\alpha_1, \dots, \alpha_{p-1})$ and the resultant length R_n . Then $(\alpha_1, \dots, \alpha_{p-1})$ and R_n are independent, viz.

$$w_{p,n}^\circ(\alpha_1, \dots, \alpha_{p-1}, r) = q_{p,n}(r) a_p^{-1} \prod_{j=2}^{p-1} \sin^{p-j} \alpha_{j-1}, \quad \forall \alpha_1, \dots, \alpha_{p-2} \in [0, \pi], \alpha_{p-1} \in [0, 2\pi), r \in (0, n],$$

iff the sample values are isotropic.

Let us denote $\underline{v} \in \mathbb{R}^p$ as $\underline{v} = u\underline{z}$, where $u = \|\underline{v}\|$ and $\underline{z} \in \mathbb{S}^{p-1}$. The moment generating function of \underline{X}_1 is given by

$$M_1(\underline{v}) = \mathbb{E}[\exp\{\langle \underline{v}, \underline{X}_1 \rangle\}] = \int_{\mathbb{S}^{p-1}} \exp\{\langle \underline{v}, \underline{x} \rangle\} dU_p(\underline{x}) = c_p^{-1}(u),$$

where \mathbb{E} denotes the expectation under isotropy. By using the recurrence relation

$$uI_{\nu+1}(u) = uI'_\nu(u) - \nu I_\nu(u), \quad (9)$$

see Abramowitz and Stegun (1965), p. 376, and by using (6), one obtains

$$\frac{d}{du} \frac{1}{c_p(u)} = \Gamma\left(\frac{u}{2}\right) \left(\frac{u}{2}\right)^{1-\frac{p}{2}} I_{\frac{p}{2}}(u).$$

From this result (and from $\partial\|\underline{v}\|/\partial\underline{v} = \underline{z}$) follows

$$\frac{\partial}{\partial\underline{v}} M_1(\underline{v}) = \frac{\partial}{\partial\underline{v}} \frac{1}{c_p(\|\underline{v}\|)} = \Gamma\left(\frac{p}{2}\right) \left(\frac{u}{2}\right)^{1-\frac{p}{2}} I_{\frac{p}{2}}(u) \underline{z} = c_p^{-1}(u) A_p(u) \underline{z}. \quad (10)$$

Denote as $K_n = \log M_n$ the cumulant generating function of \underline{R}_n . From (6) follows

$$K_n(\underline{v}) = n \left\{ \log \Gamma\left(\frac{p}{2}\right) + \log I_{\frac{p}{2}-1}(u) - \left(\frac{p}{2} - 1\right) (\log u - \log 2) \right\}$$

and from (10) follows

$$\frac{\partial}{\partial\underline{v}} K_n(\underline{v}) = n \frac{\frac{\partial}{\partial\underline{v}} M_1(\underline{v})}{M_1(\underline{v})} = n A_p(u) \underline{z}. \quad (11)$$

Let $\underline{y} \in B_n^p = \{\underline{x} \in \mathbb{R}^p \mid \|\underline{x}\| \leq n\}$, the closed ball of radius n centered at the origin, and $r = \|\underline{y}\| \in [0, n]$. The saddlepoint at $\underline{y} = r\underline{x}$ is the solution in \underline{v} of

$$\frac{\partial}{\partial\underline{v}} K_n(\underline{v}) = \underline{y} \iff n A_p(u) \underline{z} = r \underline{x} \iff A_p(u) = \frac{r}{n}, \quad \underline{z} = \underline{x}. \quad (12)$$

Denoting $\bar{\underline{v}} = \bar{u} \bar{\underline{z}}$ the saddlepoint, where $\|\bar{\underline{z}}\| = 1$, we have from (12) that

$$\bar{u} = A_p^{(-1)}\left(\frac{r}{n}\right) \quad (13)$$

and $\bar{\underline{z}} = \underline{x}$. From (8) we find

$$\bar{u} \simeq \frac{p\{p(n-r) + 4r\} - 3r - n}{8(n-r)}, \quad (14)$$

providing an approximate solution to (13) for large values of r . From (36), given below, we find

$$\bar{u} = \frac{pr}{n} + O(\bar{u}^3), \quad \text{as } r \rightarrow 0, \quad (15)$$

providing an approximate solution to (13) for small values of r . For the numerical accuracy of (14) and (15), refer to Section 4. Note also that Banerjee et al. (2005) suggest an approximation to $A_p^{(-1)}$ derived from a continued fraction, while Hornik and Grün (2014) provide upper and lower bounds to $A_p^{(-1)}$. Note that at point $\underline{y} = \mathbb{E}[\underline{X}] = 0$, both \underline{x} and \underline{z} are undetermined, whereas $r = \|\mathbb{E}[\underline{X}]\| = 0$ yields $\bar{u} = 0$.

Proposition 3.2. *The determinant of the Hessian of K_1 at $\underline{v} = u\underline{z} \in \mathbb{R}^p \setminus \{0\}$ is given by*

$$\sigma_p^2(u) = A'_p(u) \left(\frac{A_p(u)}{u} \right)^{p-1}, \quad \forall u > 0, \quad (16)$$

or alternatively by

$$\sigma_p^2(u) = \left\{ 1 - A_p^2(u) - \frac{p-1}{u} A_p(u) \right\} \left(\frac{A_p(u)}{u} \right)^{p-1}, \quad \forall u > 0.$$

Moreover

$$\lim_{u \rightarrow 0} \sigma_p^2(u) = p^{-p}. \quad (17)$$

Proof

Starting from (11), we can obtain the Hessian matrix of K_1 at $\underline{v} = u\underline{z}$ as

$$H_p(\underline{v}) = \frac{\partial}{\partial \underline{v}^\top \partial \underline{v}} K_1(\underline{v}) = \frac{A_p(u)}{u} \left\{ \left(\frac{uA'_p(u)}{A_p(u)} - 1 \right) \frac{vv^\top}{u^2} + I_{p \times p} \right\},$$

where $I_{p \times p}$ is the identity matrix of size $p \times p$. To obtain the determinant, we make use of Sylvester's determinant identity: $\det(I_{m \times m} + A_{m \times n} B_{n \times m}) = \det(I_{n \times n} + B_{n \times m} A_{m \times n})$. Thus we obtain

$$\begin{aligned} \sigma_p^2(u) &= \det H_p(u) = \left(\frac{A_p(u)}{u} \right)^p \det \left(\left\{ \frac{uA'_p(u)}{A_p(u)} - 1 \right\} \frac{vv^\top}{u^2} + I_{p \times p} \right) \\ &= \left(\frac{A_p(u)}{u} \right)^p \det \left(1 + \left\{ \frac{uA'_p(u)}{A_p(u)} - 1 \right\} \frac{v^\top v}{u^2} \right) = \frac{A'_p(u) A_p^{p-1}(u)}{u^{p-1}}. \end{aligned}$$

We now show the second formula of the proposition. From (9) and

$$I'_\nu(u) = I_{\nu-1}(u) - \frac{\nu}{u} I_\nu(u),$$

see Abramowitz and Stegun (1965), p. 376, follows that A'_p can be computed through Riccati's differential equation

$$A'_p(u) + A_p^2(u) + \frac{p-1}{u} A_p(u) = 1, \quad \forall u > 0. \quad (18)$$

The second formula is simply the first one where A'_p has been eliminated with by (18).

The third formula, i.e. the limit, can be obtained as follows. From (18) follows $\lim_{u \rightarrow 0} A'_p(u) = 1 - (p-1) \lim_{u \rightarrow 0} A_p(u)/u$. This result and l'Hôpital rule yield

$$\lim_{u \rightarrow 0} \frac{A_p(u)}{u} = \lim_{u \rightarrow 0} A'_p(u) = \frac{1}{p} \quad (19)$$

and therefore

$$\lim_{u \rightarrow 0} \sigma_p^2(u) = \lim_{u \rightarrow 0} A'_p(u) \left(\lim_{u \rightarrow 0} \frac{A_p(u)}{u} \right)^{p-1} = \frac{1}{p} \left(\frac{1}{p} \right)^{p-1}.$$

The convex conjugate or Legendre-Fenchel transform of K_1 at \underline{y}/n , for $\underline{y} \in B_n^p$, is given by

$$L(K_1) \left(\frac{\underline{y}}{n} \right) = \sup_{\underline{v} \in \mathbb{R}^p} \left\langle \underline{v}, \frac{\underline{y}}{n} \right\rangle - K_1(\underline{v}) = \bar{u} \frac{r}{n} \langle \underline{z}, \underline{x} \rangle - K_1(\bar{u} \underline{z}) = \log c_p(\bar{u}) + \bar{u} \frac{r}{n}.$$

As this convex conjugate depends on \underline{y} only through $r = \|\underline{y}\|$, it becomes practical to have the alternative notation $l(K_1)(r/n) = L(K_1)(\underline{y}/n)$. Thus the saddlepoint approximation to $g_{p,n}(\underline{y})$, the density of \underline{R}_n w.r.t. l_p at \underline{y} , is given by

$$\tilde{g}_{p,n}(\underline{y}) = (2\pi n)^{-\frac{p}{2}} \sigma_p^{-1}(\bar{u}) e^{-nl(K_1)(\frac{r}{n})} = (2\pi n)^{-\frac{p}{2}} \sigma_p^{-1}(\bar{u}) c_p^{-n}(\bar{u}) e^{-\bar{u}r}, \quad (20)$$

where

$$g_{p,n}(\underline{y}) = \tilde{g}_{p,n}(\underline{y}) \{1 + n^{-1} \xi_4(\underline{y}) + n^{-2} \xi_6(\underline{y}) + \dots\}, \quad \forall \underline{y} \in B_n^p, \quad \text{as } n \rightarrow \infty, \quad (21)$$

where the function $\xi_j(\underline{y})$ depends on the partial derivatives of K_1 up order j , evaluated at the saddlepoint at \underline{y} , for $j = 4, 6, \dots$, refer e.g. Section 6.5 of Barndorff-Nielsen and Cox (1989). At $\underline{y} = \mathbf{E}[\underline{X}] = 0$, we obtain $\tilde{g}_{p,n}(0) = \{p/(2\pi n)\}^{p/2}$. The saddlepoint approximation to the marginal density of R_n can be obtained by integrating (20) and using (4) as follows,

$$\begin{aligned} \tilde{q}_{p,n}(r) &= \int_{\mathbb{S}_r^{p-1}} \tilde{g}_{p,n}(\underline{y}) d\lambda_{p,r}(\underline{y}) \\ &= \int_{\mathbb{S}_r^{p-1}} d\lambda_{p,r}(\underline{y}) (2\pi n)^{-\frac{p}{2}} \sigma_p^{-1}(\bar{u}) c_p^{-n}(\bar{u}) e^{-\bar{u}r} \\ &= a_{p,r} (2\pi n)^{-\frac{p}{2}} \sigma_p^{-1}(\bar{u}) c_p^{-n}(\bar{u}) e^{-\bar{u}r} \\ &= 2^{1-\frac{p}{2}} \Gamma^{-1} \left(\frac{p}{2} \right) n^{-\frac{p}{2}} \sigma_p^{-1}(\bar{u}) c_p^{-n}(\bar{u}) r^{p-1} e^{-\bar{u}r}, \quad \forall r \in (0, n]. \end{aligned} \quad (22)$$

When $p = 2$, for example, we find $n^{-1} \sigma_2^{-1}(\bar{u}) I_0^n(\bar{u}) r e^{-\bar{u}r}$, as given at p. 163 of Jensen (1995) or in Proposition 2.2 of Gatto (2017). In (22) we can use

$$\sigma_p^2(\bar{u}) = \left\{ 1 - \left(\frac{r}{n} \right)^2 - \frac{(p-1)r}{n\bar{u}} \right\} \left(\frac{r}{n\bar{u}} \right)^{p-1}, \quad (23)$$

which follows from (13) and Proposition 3.2. Unless one could show that $\xi_4(\underline{y}), \xi_6(\underline{y}), \dots$, given in (21), would depend on \underline{y} only through r , the marginalization (22) does not necessarily yield a large deviations approximation, with small relative error. The large deviations nature of (22) can however be established with the Factorization lemma 3.1, which, together with Proposition 3.2, lead to the following result.

Theorem 3.3. *The saddlepoint approximation to $q_{p,n}(r)$, the density of R_n under isotropy given by (22), satisfies*

$$q_{p,n}(r) = \tilde{q}_{p,n}(r) \{1 + O(n^{-1})\}, \quad \text{as } n \rightarrow \infty,$$

where $\tilde{q}_{p,n}(r)$ is given by (22), the saddlepoint \bar{u} is the continuous increasing function of r from $[0, \infty)$ onto $[0, 1)$ given by (13), σ_p is given in Proposition 3.2 viz. (23) and c_p is given by (6), $\forall r \in (0, n]$.

Let $b_{p,n}^{-1} = \int_0^n \tilde{q}_{p,n}(r) dr$, then $b_{p,n} \tilde{q}_{p,n}(r)$ is the normalized saddlepoint approximation to $q_{p,n}(r)$ and it possesses relative error $O(n^{-3/2})$, when r is restricted to the normal deviations region.

Proof

Let us denote by $w_{p,n}^\circ(\alpha_1, \dots, \alpha_{p-1}, r)$ the joint Lebesgue density of (\underline{M}_n, R_n) in terms of the polar coordinates (1), under isotropy. Then from (2), (3), (20) and (21) follows

$$\begin{aligned} w_{p,n}^\circ(\alpha_1, \dots, \alpha_{p-1}, r) &= g_{p,n}(\underline{g}(\alpha_1, \dots, \alpha_{p-1}, r)) J(\alpha_1, \dots, \alpha_{p-2}, r) \\ &= \left[(2\pi n)^{-\frac{p}{2}} \sigma_p^{-1}(\bar{u}) c_p^{-n}(\bar{u}) e^{-\bar{u}r} \{1 + O(n^{-1})\} r^{p-1} a_{p,1} \right] \left[a_{p,1}^{-1} \prod_{j=2}^{p-1} \sin^{p-j} \alpha_{j-1} \right] \\ &= [\tilde{q}_{p,n}(r) \{1 + O(n^{-1})\}] \left[a_p^{-1} \prod_{j=2}^{p-1} \sin^{p-j} \alpha_{j-1} \right], \quad \forall r \in (0, n]. \end{aligned}$$

From the Factorization lemma 3.1 follows that the first squared bracket must contain the marginal density of R_n , because the second squared bracket contains the polar form of the isotropic density over \mathbb{S}^{p-1} .

For a justification of the $O(n^{-3/2})$ relative error over the normal deviations region after normalization, see e.g. p. 31 of Field and Ronchetti (1990). If, however, $\xi_4(\underline{y})$ given in (21) were independent of \underline{y} , then the relative error after normalization would be $O(n^{-2})$ over the large deviations region. \square

Note that the exact density of the resultant length under isotropy can be obtained by inversion of its characteristic function, refer e.g. to Section 15.4.1 Mardia et al. (1979) for explanations. Its density at $r \in [0, n]$ is $r\psi_n(r)$, where

$$\psi_n(r) = 2^{(1-\frac{p}{2})(n-1)} \Gamma^{n-1} \left(\frac{p}{2} \right) \int_0^\infty t^{n-\frac{1}{2}(n-1)p} J_{\frac{p}{2}-1}(rt) J_{\frac{p}{2}-1}^n(t) dt \quad (24)$$

and where J_ν denotes the Bessel function of the first kind of order ν , which can be obtained through

$$J_\nu(z) = e^{\frac{\nu}{2}\pi i} I_\nu(e^{-\frac{\nu}{2}\pi i} z), \quad \text{for } -\pi < \arg z < \frac{\pi}{2} \quad \text{and } \Re \nu > -\frac{1}{2},$$

where I_ν is given in (7), cf. Abramowitz and Stegun (1965), p. 375. Thus the saddlepoint approximation (22) provides an approximation to ψ_n . Note that the infinite integration domain and the fast oscillating integrand in (24) make numerical integration difficult.

We turn to approximations to upper tail probabilities $P[R_n \geq r]$, with $r \in [0, n]$ possibly large. The first approach is the numerical integration of the (smooth) saddlepoint approximation to the density, after the change of integration variable which avoids the necessity of computing the saddlepoint at each ordinate. Let $r \in (0, n)$ and \bar{u} defined by (13). Then, by

using (22) and (16), we obtain

$$\begin{aligned}\tilde{S}_{p,n}(r) &= b_{p,n} \int_r^n \tilde{q}_{p,n}(s) ds = b_{p,n} \int_{\bar{u}}^\infty \tilde{q}_{p,n}(s(u)) \frac{ds(u)}{du} du \\ &= b_{p,n} \frac{n^{\frac{p}{2}}}{2^{\frac{p}{2}-1} \Gamma\left(\frac{p}{2}\right)} \int_{\bar{u}}^\infty u^{p-1} \sigma_p(u) \exp\{-n[uA_p(u) + \log c_p(u)]\} du\end{aligned}\quad (25)$$

as an approximation to $P[R_n \geq r]$. The normalizing constant $b_{p,n}$ can be evaluated with the above integral with $\bar{u} = 0$. This formula is simple and fast to compute and \bar{u} can be computed by taking (14) or (15) as initial values, for upper or lower tails respectively.

The second approach applies the asymptotic expansion of the integral of the saddlepoint approximation to the density of Temme (1982). This leads to a simple formula, which is accurate for small upper tail probabilities only. This solution was suggested by Jensen (1995), p. 164-165, but for the isotropic random walk over the plane i.e. for $p = 2$ only. The generalization to arbitrary dimension p is the following.

Theorem 3.4. *An asymptotic approximation to the upper tail probability of R_n under isotropy, viz. $P[R_n \geq r]$, is given by*

$$\tilde{S}_{p,n}^\dagger(r) = \frac{\sqrt{2^{3-p}\pi}}{\Gamma\left(\frac{p}{2}\right) n^{\frac{p}{2}} \bar{u} \sigma_p(\bar{u})} r^{p-1} \phi(\sqrt{n} s_n(r)), \quad (26)$$

and

$$\tilde{S}_{p,n}(r) = \tilde{S}_{p,n}^\dagger(r) \{1 + O(n^{-2})\}, \quad \text{as } n \rightarrow \infty,$$

where ϕ denotes the standard normal density,

$$s_n(r) = \sqrt{2l(K_1)\left(\frac{r}{n}\right)} = \sqrt{2\left\{\frac{r}{n}\bar{u} + \log c_p(\bar{u})\right\}}, \quad (27)$$

\bar{u} is the saddlepoint given by (13), σ_p is given in Proposition 3.2 viz. (23) and c_p is given by (6), $\forall r \in (0, n]$.

Proof

Let $r \in [0, n)$. Integrating $\tilde{q}_{p,n}$ given in (22) with a change of variables yields

$$b_{p,n}^{-1} \tilde{S}_{p,n}(r) = \int_r^n \tilde{q}_{p,n}(x) dx = \frac{\sqrt{2^{3-p}\pi}}{\Gamma\left(\frac{p}{2}\right) n^{\frac{p+1}{2}}} \int_{s(r)}^\infty \left(\frac{n}{2\pi}\right)^{\frac{1}{2}} \exp\left\{-\frac{n}{2}s^2\right\} h_n(s) ds, \quad (28)$$

where s is equal to s_n given by (27),

$$h_n(s) = \frac{x^{p-1}(s) \frac{dx(s)}{ds}}{\sigma_p(\bar{u})} \quad (29)$$

and \bar{u} is given by (13) with $x(s)$ replacing r . One obtains

$$\frac{ds(x)}{dx} = s^{-1}(x) \frac{d}{dx} \left\{ \frac{x}{n} \bar{u} + \log c_p(\bar{u}) \right\} = \frac{\bar{u}}{s(x)n}. \quad (30)$$

Indeed, from (11) follows $\partial/(\partial \underline{v}) \log c_p(\|\underline{v}\|) = -A_p(\|\underline{v}\|)\underline{v}/\|\underline{v}\|$, which simplifies to

$$\frac{c'_p(u)}{c_p(u)} = -A_p(u). \quad (31)$$

Thus $c'_p(\bar{u})/c_p(\bar{u}) = -x/n$, which justifies (30). We can now apply the asymptotic expansion for integrals of Temme (1982), reported at p. 67 of Jensen (1995), and obtain that the integral in (28) is equal to

$$\left\{1 - \Phi(\sqrt{n}s(r))\right\} \frac{1}{n} \frac{h''_n(0)}{2} + \frac{1}{\sqrt{n}} \phi(\sqrt{n}s(r)) \frac{h_n(s(r)) - h_n(0)}{s(r)} + O\left(n^{-\frac{3}{2}} \phi(\sqrt{n}s(r))\right), \quad (32)$$

as $n \rightarrow \infty$, where Φ denotes the standard normal distribution function. From (13), (17) and (19) we deduce $\lim_{s \rightarrow 0} h_n(s) = 0$. In the next paragraph, we show that $\lim_{s \rightarrow 0} h''_n(s)$ is negligible. Thus, omitting $h_n(0) = h''_n(0)$ in (32) and simplifying, considering

$$1 - \Phi(z) = \frac{\phi(z)}{z} \{1 + O(z^{-2})\}, \quad \text{as } z \rightarrow \infty, \quad (33)$$

leads to the desired result.

We now show by series expansions that $\lim_{s \rightarrow 0} h''_n(s)$ is negligible. Define $\tilde{h}_n = (n^{p-2}h_n)^2/2$. From (16) and (30), we find

$$\tilde{h}_n = \frac{\{A_p(\bar{u})\bar{u}\}^{p-1} \{A_p(\bar{u})\bar{u} + \log c_p(\bar{u})\}}{A'_p(\bar{u})\bar{u}^2}.$$

We now note that from (18) follows

$$A''_p(u) = -2A_p(u)A'_p(u) - \frac{p-1}{u} \left\{ A'_p(u) - A'_p(0) - \frac{u}{2}A''_p(0) \right\} + o(1), \quad \text{as } u \rightarrow 0.$$

Letting $u \rightarrow 0$ yields $A''_p(0) = -(p-1)A''_p(0)/2$ and thus

$$A''_p(0) = 0. \quad (34)$$

From (19), (31) and (34) follows

$$\log c_p(u) = -uA_p(0) - \frac{u^2}{2}A'_p(0) - \frac{u^3}{6}A''_p(0) + O(u^4) = -\frac{u^2}{2p} + O(u^4), \quad \text{as } u \rightarrow 0. \quad (35)$$

From this result and from

$$A_p(u)u = \frac{u^2}{p} + O(u^4), \quad \text{as } u \rightarrow 0, \quad (36)$$

follows

$$\tilde{h}_n = \frac{1}{2p^{p-1}} \frac{\bar{u}^{2p-2} + O(\bar{u}^{2p})}{1 + O(\bar{u}^2)} = \frac{1}{2p^{p-1}} \bar{u}^{2(p-1)} \{1 + O(\bar{u}^2)\}, \quad \text{as } s \rightarrow 0.$$

So we obtain

$$h_n = n^{2-p} p^{\frac{1-p}{2}} \bar{u}^{p-1} \{1 + O(\bar{u}^2)\}, \quad \text{as } s \rightarrow 0. \quad (37)$$

Two simple results are

$$\frac{d\bar{u}}{dx} = \frac{1}{A'_p(\bar{u})n} \rightarrow \frac{p}{n}, \quad \text{as } s \rightarrow 0, \quad (38)$$

and

$$\lim_{s \rightarrow 0} \frac{dx}{ds} = \frac{n}{\sqrt{p}}, \quad (39)$$

from (19) and by noting $l = \lim_{s \rightarrow 0} dx/ds = \lim_{s \rightarrow 0} ns/\bar{u} = n \lim_{s \rightarrow 0} s/x \lim_{s \rightarrow 0} x/\bar{u} = n^2(lp)^{-1}$. Consequently, given (30), one finds

$$\frac{d\bar{u}}{ds} = \frac{d\bar{u}}{dx} \frac{dx}{ds} = \frac{s}{\bar{u}A'_p(\bar{u})} \rightarrow \frac{p}{n} \frac{n}{\sqrt{p}} = \sqrt{p}, \quad \text{as } s \rightarrow 0, \quad (40)$$

and

$$\frac{d^2\bar{u}}{ds^2} = \frac{d}{ds} \frac{s}{\bar{u}A'_p(\bar{u})} = \frac{1}{\bar{u}A'_p(\bar{u})} \left\{ 1 - \frac{s^2}{\bar{u}^2 A'_p(\bar{u})} - \frac{s^2 A''_p(\bar{u})}{\bar{u} [A'_p(\bar{u})]^2} \right\} \quad (41)$$

$$= \frac{1}{\bar{u}A'_p(\bar{u})} \left\{ 1 - \underbrace{\frac{s}{\bar{u}} \frac{d\bar{u}}{ds}}_{=1+o(\bar{u})} - \underbrace{\frac{s}{\bar{u}} \frac{d\bar{u}}{ds}}_{=1+o(\bar{u})} \underbrace{\frac{\bar{u}A''_p(\bar{u})}{A'_p(\bar{u})}}_{=o(\bar{u})} \right\} = o(1), \quad \text{as } s \rightarrow 0, \quad (42)$$

from (19) and (34). With (37), (40) and (41) we finally obtain, for $p = 3, 4, \dots$,

$$\begin{aligned} h''_n(s) &= \frac{d^2 h_n}{d\bar{u}^2} \left(\frac{d\bar{u}}{ds} \right)^2 + \frac{dh_n}{d\bar{u}} \frac{d^2 \bar{u}}{ds^2} = n^{2-p} p^{1-\frac{p}{2}} (p-1)(p-2) \bar{u}^{p-3} + o(1) \\ &\rightarrow \begin{cases} 0, & \text{if } p = 4, 5, \dots, \\ \frac{2}{n}, & \text{if } p = 3, \end{cases} \quad \text{as } s \rightarrow 0. \end{aligned}$$

For the case $p = 3$, (33) implies that $h''_n(0)$ has no contribution in (32) and can be ignored. For the case $p = 2$, $h''_n(s) \rightarrow 0$, as $s \rightarrow 0$, can be obtained in the same way. \square

As seen in Section 4, the approximation of Theorem 3.4 is most accurate for small upper tail probabilities or with large sample sizes. This is due to the form of the error term in (32). For larger upper tail probabilities or for very small sample sizes, the approximation (25) should be preferred.

A consequence of (35) is the small concentration approximation to the normalizing constant of the von Mises-Fisher distribution

$$c_p(\kappa) = \exp \left\{ -\frac{1}{2} \frac{\kappa^2}{p} \right\} \{1 + O(\kappa^4)\}, \quad \text{as } \kappa \rightarrow 0. \quad (43)$$

3.2 Von Mises-Fisher random walk

Consider now the vMF($\underline{\mu}, \kappa$) distribution with density (5) for the independent sample directions $\underline{X}_1, \dots, \underline{X}_n$. Denote by $q_{\kappa,p,n}$ the Lebesgue density of the resultant length R_n , under

the $\text{vMF}(\underline{\mu}, \kappa)$ distribution. Thus $q_{0,p,n} = q_{p,n}$. The joint density of (\underline{M}_n, R_n) w.r.t. $U_p \times l_1$ and at $(\underline{x}, r) \in \mathbb{S}^{p-1} \times (0, n]$ is given by

$$\begin{aligned} w_{\kappa,p,n}(\underline{x}, r) &= \int \dots \int_{x_1 + \dots + x_n = rx} c_p^n(\kappa) \exp\{\kappa \langle \underline{\mu}, \underline{x}_1 + \dots + \underline{x}_n \rangle\} dU_p(\underline{x}_1) \dots dU_p(\underline{x}_n) \\ &= c_p^n(\kappa) \exp\{\kappa r \langle \underline{\mu}, \underline{x} \rangle\} w_{0,p,n}(\underline{x}, r) = c_p^n(\kappa) \exp\{\kappa r \langle \underline{\mu}, \underline{x} \rangle\} q_{0,p,n}(r), \end{aligned} \quad (44)$$

where the last equality follows from the Factorization lemma 3.1 and where $w_{0,p,n} = w_{p,n}$. The density of R_n at $r \in (0, n]$ is given by $q_{\kappa,p,n}(r) = \int_{\mathbb{S}^{p-1}} w_{\kappa,p,n}(\underline{x}, r) dU_p(\underline{x})$. This and (44) justify the following result.

Lemma 3.5 (Tilting lemma). *The densities of R_n w.r.t. l_1 , under isotropy and under the $\text{vMF}(\underline{\mu}, \kappa)$ distribution, respectively denoted $q_{p,n}$ and $q_{\kappa,p,n}$, satisfy the relation*

$$q_{\kappa,p,n}(r) = \frac{c_p^n(\kappa)}{c_p(\kappa r)} q_{p,n}(r), \quad \forall r \in (0, n].$$

Theorem 3.3 and Lemma 3.5 lead directly to the following result.

Theorem 3.6. *The saddlepoint approximation to $q_{\kappa,p,n}(r)$, the density of R_n w.r.t. l_1 under the $\text{vMF}(\underline{\mu}, \kappa)$ distribution, is given by*

$$\tilde{q}_{\kappa,p,n}(r) = 2^{1-\frac{p}{2}} \Gamma^{-1}\left(\frac{p}{2}\right) n^{-\frac{p}{2}} \sigma_p^{-1}(\bar{u}) \left(\frac{c_p(\kappa)}{c_p(\bar{u})}\right)^n c_p^{-1}(\kappa r) r^{p-1} e^{-\bar{u}r} \quad (45)$$

and satisfies

$$q_{\kappa,p,n}(r) = \tilde{q}_{\kappa,p,n}(r) \{1 + O(n^{-1})\}, \quad \text{as } n \rightarrow \infty,$$

where the saddlepoint \bar{u} is the continuous increasing function of r from $[0, \infty)$ onto $[0, 1)$ given by (13), σ_p is given in Proposition 3.2 viz. (23) and c_p is given by (6), $\forall r \in (0, n]$.

Let $b_{\kappa,p,n}^{-1} = \int_0^n \tilde{q}_{\kappa,p,n}(r) dr$, then $b_{\kappa,p,n} \tilde{q}_{\kappa,p,n}(r)$ is the normalized saddlepoint approximation to $q_{\kappa,p,n}(r)$ and it possesses relative error $O(n^{-3/2})$, when r is restricted to the normal deviations region.

Note that, as expected, both the density $q_{\kappa,p,n}$ and its saddlepoint approximation $\tilde{q}_{\kappa,p,n}$ do not depend on $\underline{\mu}$.

Let \mathbf{P}_κ denote the probability measure under the $\text{vMF}(\underline{\mu}, \kappa)$ distribution and, thus, $\mathbf{P}_0 = \mathbf{P}$. We interested in tail probabilities of R_n under the $\text{vMF}(\underline{\mu}, \kappa)$ distribution, viz. in $\mathbf{P}_\kappa[R_n \geq r]$, for $r \in [0, n]$. The approach using numerical integration of the saddlepoint approximation to the density is straightforward. Let $r \in (0, n)$ and \bar{u} defined by (13). Then from (45) follows

$$\tilde{S}_{\kappa,p,n}(r) = b_{\kappa,p,n} \frac{n^{\frac{p}{2}} c_p^n(\kappa)}{2^{\frac{p}{2}-1} \Gamma\left(\frac{p}{2}\right)} \int_{\bar{u}}^{\infty} \frac{u^{p-1} \sigma_p(u)}{c_p(n\kappa A_p(u))} \exp\{-n[uA_p(u) + \log c_p(u)]\} du$$

as an approximation to $\mathbf{P}_\kappa[R_n \geq r]$ and we clearly have $\tilde{S}_{0,p,n}(r) = \tilde{S}_{p,n}(r)$, given that $b_{0,p,n} = b_{p,n}$.

The second approach with the asymptotic expansion of Temme (1982) yields the following generalization of Theorem 3.4 to the random walk with von Mises-Fisher directions, for r large.

Theorem 3.7. *An asymptotic approximation to the upper tail probability of R_n under the vMF($\underline{\mu}, \kappa$) distribution, viz. $P_\kappa[R_n \geq r]$, is given by*

$$\tilde{S}_{\kappa,p,n}^\dagger(r) = \frac{c_p^n(\kappa)}{c_p(\kappa r)} \tilde{S}_{p,n}^\dagger(r)$$

and

$$\tilde{S}_{\kappa,p,n}(r) = \tilde{S}_{\kappa,p,n}^\dagger(r) \{1 + O(n^{-2})\}, \quad \text{as } n \rightarrow \infty,$$

where $\tilde{S}_{p,n}^\dagger(r) = \tilde{S}_{0,p,n}^\dagger(r)$ is given in Theorem 3.4, $\forall r \in (0, n]$.

This theorem can re-expressed as follows. The saddlepoint approximation to $P_\kappa[R_n \geq r]$ is provided by $\tilde{S}_{\kappa,p,n}^\dagger(r)$, which is a simple transform of the saddlepoint approximation under isotropy $\tilde{S}_{p,n}^\dagger(r)$. The saddlepoint approximation $\tilde{S}_{\kappa,p,n}^\dagger(r)$ possesses a small relative error of the order n^{-2} w.r.t. $\tilde{S}_{\kappa,p,n}(r)$, viz. the approximation obtained by integrating the saddlepoint approximation to the density. As this integral must be evaluated numerically, $\tilde{S}_{\kappa,p,n}^\dagger(r)$ should be faster to compute.

Proof

This theorem is the consequence of the application of the Tilting lemma 3.5 to Theorem 3.4. Precisely, define

$$h_{\kappa,n}(s) = \frac{h_n(s)}{c_p(\kappa x(s))},$$

where $h_n(s)$ is given by (29). If $h_n(0) = 0$, then $h_{\kappa,n}(0) = 0$. From (37) and (43) follows

$$h_{\kappa,n} = n^{2-p} p^{\frac{1-p}{2}} \hat{h}_{\kappa,n} \{1 + O(\bar{u}^2)\}, \quad \text{as } s \rightarrow 0,$$

where

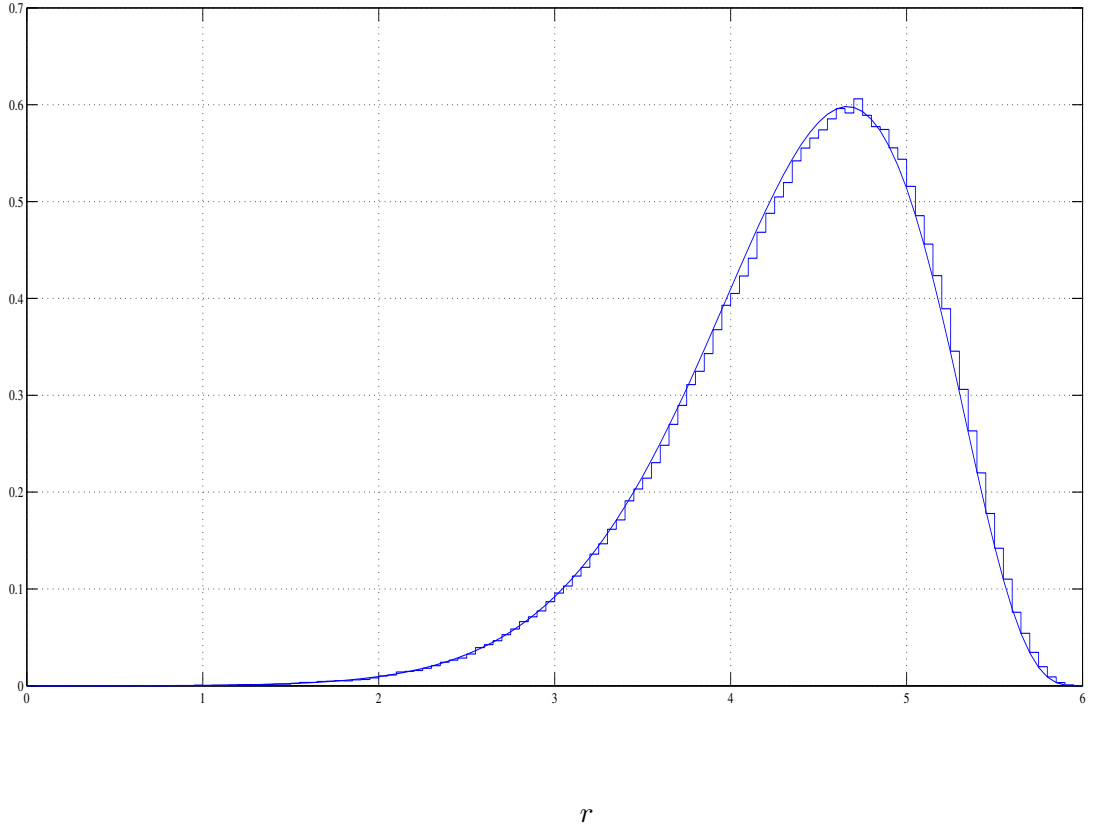
$$\hat{h}_{\kappa,n} = \bar{u}^{p-1} \exp \left\{ \frac{1}{2} \frac{(\kappa x)^2}{p} \right\}$$

and where the asymptotic order of the error can be justified by (38). For $p = 3, 4, \dots$, we have

$$\begin{aligned} \frac{d^2 \hat{h}_{\kappa,n}}{ds^2} = \exp \left\{ \frac{1}{2} \frac{(\kappa x)^2}{p} \right\} & \left[(p-1)(p-2) \bar{u}^{p-3} \left(\frac{d\bar{u}}{ds} \right)^2 + (p-1) \bar{u}^{p-2} \frac{d^2 \bar{u}}{ds^2} \right. \\ & \left. + 2(p-1) \bar{u}^{p-2} \frac{d\bar{u}}{ds} \frac{\kappa^2 x}{p} \frac{dx}{ds} + \bar{u}^{p-1} \left(\frac{\kappa^2 x}{p} \frac{dx}{ds} \right)^2 + \bar{u}^{p-1} \frac{\kappa^2}{p} \left\{ \left(\frac{dx}{ds} \right)^2 + x \frac{d^2 x}{ds^2} \right\} \right]. \quad (46) \end{aligned}$$

For $p = 4, 5, \dots$, all five main summands in the square bracket above vanish as s vanishes: because of (40) for the first summand, because of (41) for the second, because of (39) and (40) for the third, because of (39) for the fourth and because of (30) and (39) for the fifth summand, in which the curly bracket can be re-expressed as $n/\bar{u}(1 - s/\bar{u} d\bar{u}/ds)$. For $p = 3$, the only nonvanishing summand among the five main summands in the square bracket of (46) is the first one and it converges to 6, from (40). But $\lim_{s \rightarrow 0} h_{\kappa,n}''(s) = 2/n$, which brings no contribution to (32). For $p = 2$, $\lim_{s \rightarrow 0} h_{\kappa,n}''(s) = 0$ can be obtained in the same way. Thus we are simply led to replace $h_n(s(r))$ by $h_{\kappa,n}(s(r))$ in Temme's asymptotic expansion (32) and the theorem follows directly. \square

Figure 1: *Saddlepoint approximation with normalization and histogram of simulations with $p = 3$, $\kappa = 3$ and $n = 6$*



4 Numerical study

This section provides numerical illustrations of the accuracy of the saddlepoint approximations to the density and to the tail probabilities of the resultant length R_n with isotropic and von Mises-Fisher directions. It also shows a small empirical study of the behavior of the densities of R_n under various values of the concentration κ and the dimension p . This last study is made simple thanks to the computational efficiency of the saddlepoint approximation. The accuracy of the saddlepoint approximation always refers to the distribution obtained by Monte Carlo simulation, based on $5 \cdot 10^5$ generations.

We first consider $p = 3$, $\kappa = 3$ and $n = 6$. The normalizing constant of the saddlepoint approximation is $b_{3,3,6}^{-1} = 1.2127$. Figure 1 shows the histogram of the simulated resultant lengths (step function) together with the normalized saddlepoint approximation (continuous function), $b_{3,3,6}\tilde{q}_{3,3,6}$, see (45): the saddlepoint approximations appears very accurate.

We now consider $p = 4$, $\kappa = 0, 3, 8$ and $n = 6$. The normalizing constants of the considered saddlepoint approximations to these densities are $b_{0,4,6}^{-1} = 1.2143$, $b_{3,4,6}^{-1} = 1.2516$ and $b_{8,4,6}^{-1} = 1.2640$. Figure 2 shows the three normalized saddlepoint approximations with $\kappa = 0, 3, 8$ respectively from the left to the right.

We finally consider $p = 3, 5, 9$, $\kappa = 3$ and $n = 6$. The normalizing constants are $b_{3,3,6}^{-1} = 1.2127$, $b_{3,5,6}^{-1} = 1.2983$ and $b_{3,9,6}^{-1} = 1.5295$. Figure 3 shows the three normalized saddlepoint

Figure 2: *Saddlepoint approximations with normalization, with $p = 4$, $\kappa = 0, 3, 8$, $n = 6$ and with $\kappa = 0, 3, 8$, appearing respectively from the left to the right according to these values of κ*

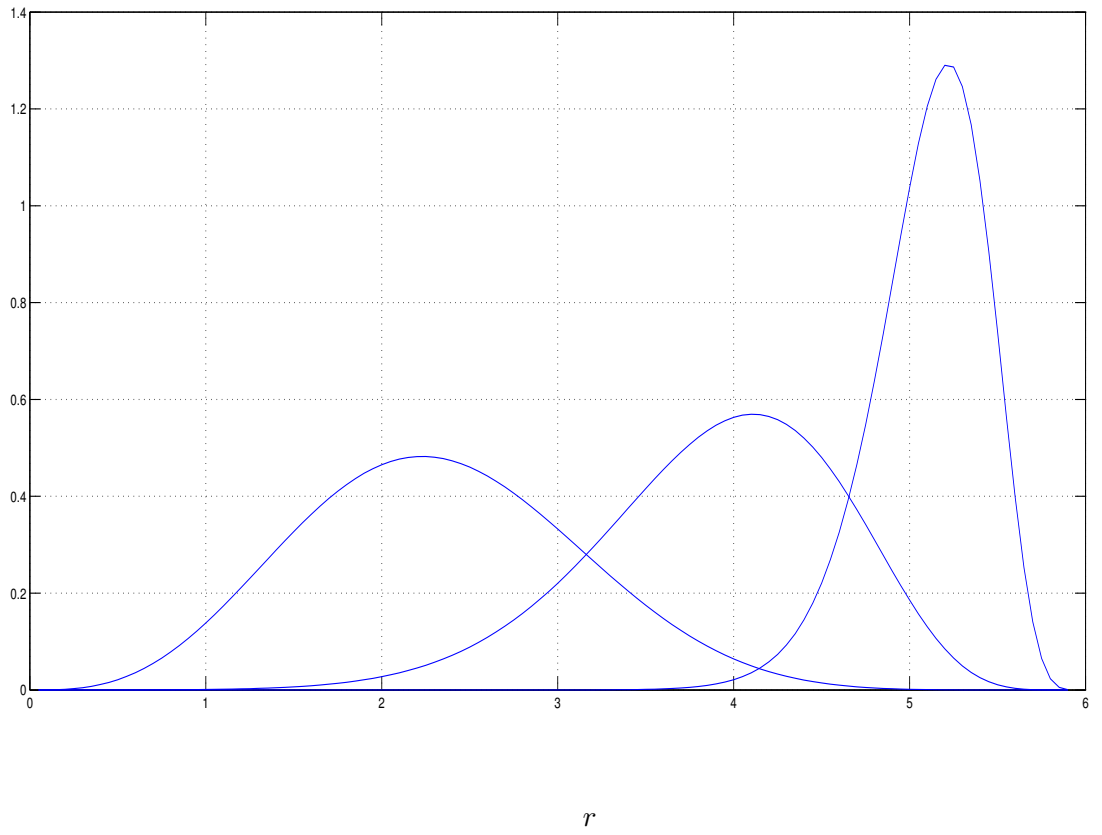
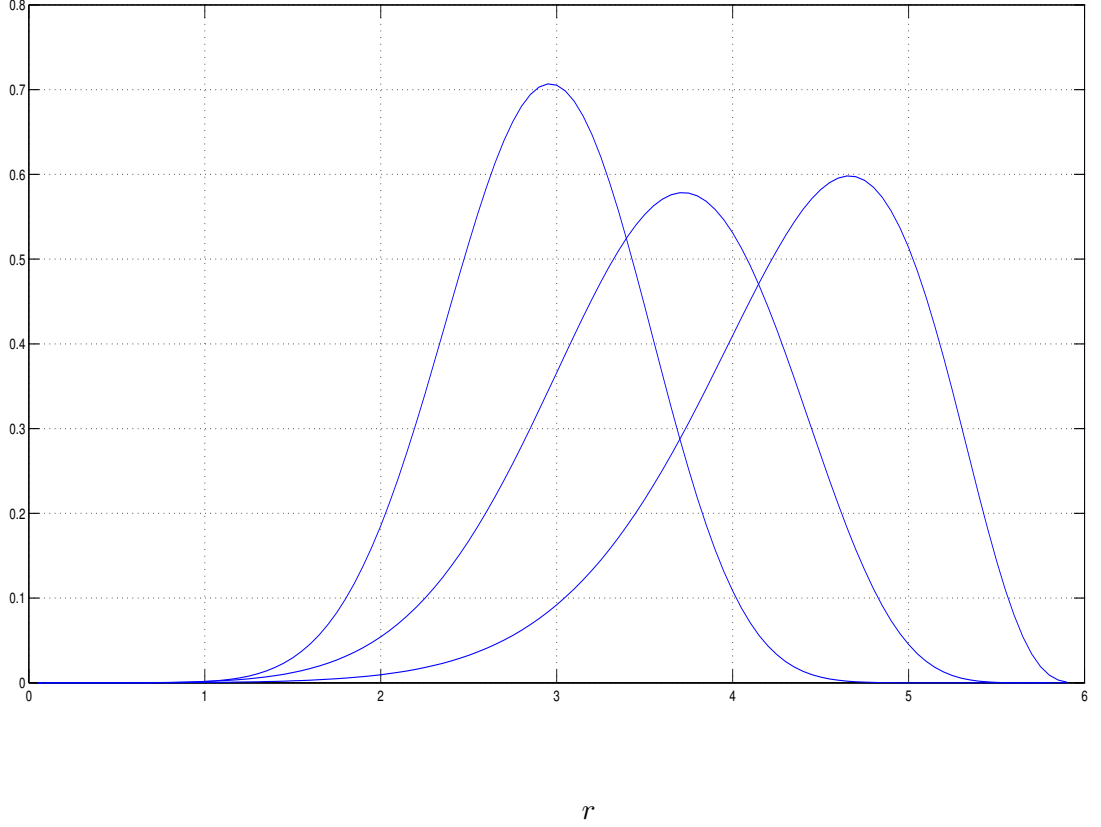


Figure 3: *Saddlepoint approximations with normalization, with $\kappa = 3$, $n = 6$ and with $p = 3, 5, 9$, appearing respectively from the right to left according to these values of p*

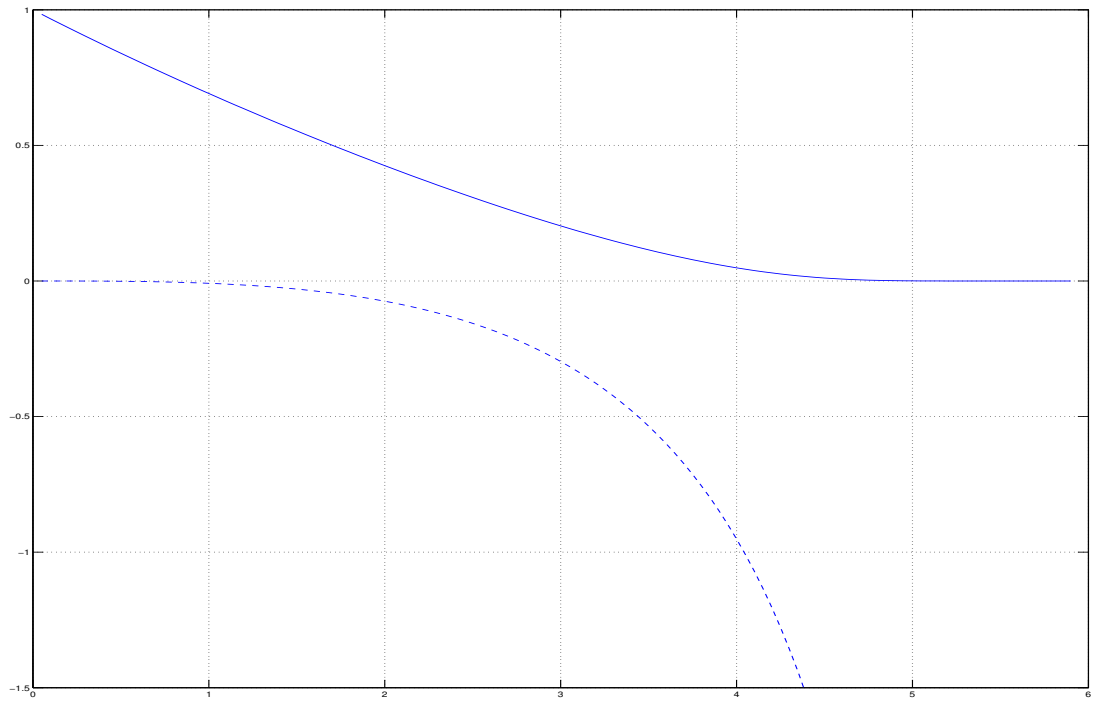


approximations with $p = 3, 5, 9$ respectively from the right to the left. Note that we obtain normalizing constants significantly larger than one because we are considering a very small sample size n , relatively to the dimension p . The following computations show that the magnitude of the normalizing constant decreases to one as the sample size n increases: $b_{3,9,10}^{-1} = 1.2783$, $b_{3,9,20}^{-1} = 1.1268$ and $b_{3,9,30}^{-1} = 1.0821$.

Figure 4 illustrates the behavior of the approximate solutions to the saddlepoint equation given by (14), for large values of r , see solid line, and (15), for vanishing values of r , see dashed line. Each one of the two curves give the approximate saddlepoint minus the saddlepoint \bar{u} obtained by solving the saddlepoint equation (13) iteratively, as explained in the last paragraph. The parameter values are $p = 3$ and $n = 6$. As we can see, these approximations are accurate with either small or large values of r , as desired.

We finally consider the isotropic random walk with $p = 3$ and $n = 10$. The normalizing constant of the saddlepoint approximation is $b_{0,3,10}^{-1} = 1.0867$. Table 1 gives the simulated upper tails probabilities $\mathbf{P}^{\text{MC}}[R_{10} \geq r]$ together with the approximation $\tilde{S}_{3,10}(r)$, obtained by integrating the normalized saddlepoint approximation to the density, see (25), and with the direct formula for tail probabilities $\tilde{S}_{3,10}^{\dagger}(r)$, see (26), for values of r between 0.16 and 0.0001. We can see that the integrated saddlepoint approximation $\tilde{S}_{3,10}$ is always very accurate. The direct approximation $\tilde{S}_{3,10}^{\dagger}(r)$ is very accurate for very small and small upper tail probability,

Figure 4: *Errors of the approximations to the saddlepoint \bar{u} given by (14) minus exact value - solid line - and (15) minus exact value - dashed line - with $p = 3$ and $n = 6$*



r

Table 1: *Saddlepoint approximations and Monte Carlo upper tail probabilities under isotropy with $p = 3$ and $n = 10$*

r	$\mathbf{P}^{\text{MC}}[R_{10} \geq r]$	$\tilde{S}_{3,10}(r)$	$\tilde{S}_{3,10}^{\dagger}(r)$
4.12	0.1652	0.1676	0.1450
4.27	0.1393	0.1413	0.1233
4.39	0.1205	0.1224	0.1075
4.54	0.0996	0.1014	0.0897
4.71	0.0796	0.0810	0.0722
4.91	0.0598	0.0611	0.0549
5.17	0.0401	0.0412	0.0373
5.57	0.0201	0.0210	0.0192
5.93	0.0101	0.0106	0.0098
6.04	0.0080	0.0085	0.0078
6.18	0.0060	0.0063	0.0058
6.35	0.0040	0.0043	0.0040
6.64	0.0020	0.0021	0.0020
6.89	0.001016	0.001103	0.001027
7.14	0.000496	0.000537	0.000500
7.56	0.000102	0.000135	0.000126

perhaps below 10%, and loses accuracy for larger tail probabilities, in accordance with the remark given at the end of Section 3.1.

This numerical study is performed with *Matlab* and the function `fzero` is used for solving the saddlepoint equation. *Matlab*'s programs used for these computations are available at <http://www.stat.unibe.ch>. *Matlab*'s program for generating von Mises-Fisher directions is part of Chen et al. (2015) and it is available at <http://www.mathworks.com/matlabcentral/fileexchange>.

5 References

- Abramowitz, M., Stegun, I. E. (1972), *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover (reprint).
- Amos, D. E. (1974), "Computation of modified Bessel functions and their ratios", *Mathematics of Computation*, 28, 239-251.
- Banerjee, A., Dhillon, I. S., Ghosh, J., Sra, S. (2005), "Clustering on the unit hypersphere using von Mises-Fisher distributions", *Journal of Machine Learning Research*, 6, 1345-1382.
- Barakat, R. (1973), "Isotropic random flights", *Journal of Physics A: Mathematical, Nuclear and General*, 6, 796-804.

- Barber, M. N., Ninham, B. W. (1970), *Random and Restricted Walks, Theory and Applications*, Gordon and Breach.
- Barndorff-Nielsen, O. E., Cox, D. R. (1989), *Asymptotic Techniques for Use in Statistics*, Chapman & Hall.
- Chen, Y.-H., Wei, D., Newstadt, G., DeGraef, M., Simmons, J., Hero, A. (2015), “Parameter Estimation in Spherical Symmetry Groups”, *IEEE Signal Processing Letters*, 22, 1152-1155.
- Daniels, H. E. (1954), “Saddlepoint approximations in statistics”, *The Annals of Mathematical Statistics*, 25, 631-650.
- Daniels, H. E., Young, G. A. (1991), “Saddlepoint approximation for the studentized mean, with an application to the bootstrap”, *Biometrika*, 78, 169-179.
- De Bruijn, N. G. (1982), *Asymptotic Methods in Analysis*, Dover (reprint).
- Feuerverger, A. (1989), “On the empirical saddlepoint approximation”, *Biometrika*, 76, 457-464.
- Field, C. (1982), “Small sample asymptotic expansions for multivariate M -estimates”, *The Annals of Statistics*, 10, 672-689.
- Field, C. A., Ronchetti, E. (1990), *Small Sample Asymptotics*, Institute of Mathematical Statistics, 13.
- Field, C. A., Tingley, M. A. (1997), “Small sample asymptotics: applications in robustness”, *Handbook of Statistics*, Elsevier, Volume 15, Chapter 18, 513-536.
- Flory, P. J. (1969), *Statistical Mechanics of Chain Molecules*, Interscience.
- Gatto, R. (2000), “Multivariate saddlepoint test for the wrapped normal model”, *Journal of Statistical Computation and Simulation*, 65, 271-285.
- Gatto, R. (2017), “Large deviations approximations to distributions of the total distance of compound random walks with von Mises directions”, *Methodology and Computing in Applied Probability*, to appear.
- Gatto, R., Jammalamadaka, S. R. (1999), “A conditional saddlepoint approximation for testing problems”, *Journal of the American Statistical Association*, 94, 533-541.
- Gatto, R., Jammalamadaka, S. R. (2003). “Inference for wrapped symmetric α -stable circular models”, *Sankhyā, A: Mathematical Statistics and Probability*, 65, 2, 333-355.
- Gatto, R., Mayer, M. (2005), “Saddlepoint approximations for some models of circular data”, *Statistical Methodology*, 2, 233-248.
- Hornik, K., Grün, B. (2014), “On maximum likelihood estimation of the concentration parameter of von Mises-Fisher distributions”, *Computational Statistics*, 29, 945-957.
- Jeganathan, P., Paige, R. L., Trindade A. A. (2015), “Saddlepoint-based bootstrap inference for the spatial dependence parameter in the lattice process”, *Spatial Statistics*, 12, 1-14.
- Jensen, J. L. (1995), *Saddlepoint Approximations*, Oxford University Press.

- Kent, J. T., Mardia, K. V., Rao, J. S. (1979), "A characterization of uniform distribution on the circle", *Annals of Statistics*, 7, 882-889.
- Kolassa, J. E. (2006), *Series Approximation Methods in Statistics*, Lecture Notes in Statistics, 88, third edition, Springer.
- Kume, A., Wood, A. T. A. (2005), "Saddlepoint Approximations for the Bingham and Fisher-Bingham normalizing constants", *Biometrika*, 92, 465-476.
- Lugannani, R., Rice, S. (1980), "Saddle point approximation for the distribution of the sum of independent random variables", *Advances in Applied probability*, 12, 475-490.
- Ma, Y., Ronchetti, E. (2011), "Saddlepoint Test in Measurement Error Models", *Journal of the American Statistical Association*, 106, 147-156.
- Mardia, K. V., Jupp, P. E. (2000), *Directional Statistics*, Wiley.
- Mardia, K. V., Kent, J. T., Bibby J. M. (1979), *Multivariate Analysis*, Academic Press.
- Masoliver, M., Porrá, J. M., Weiss, G. H. (1993), "Some two and three-dimensional persistent random walk", *Physica A: Statistical Mechanics and its Applications*, 469-482.
- Orsingher, E., De Gregorio, A. (2007), "Random flights in higher spaces", *Journal of Theoretical Probability*, 20, 769-806.
- Reid, N. (1988), "Saddlepoint methods and statistical inference", *Statistical Science*, 3, 213-238.
- Robinson, J. (1982), "Saddlepoint approximations for permutation tests and confidence intervals", *Journal of the Royal Statistical Society, B*, 44, 91-101.
- Robinson, J., Ronchetti, E., Young, G. A. (2003), "Saddlepoint approximations and tests based on multivariate M -estimates", *The Annals of Statistics*, 31, 1154-1169.
- Ronchetti, E., Welsh, A. H. (1994), "Empirical saddlepoint approximations for multivariate M -estimators", *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*, 56, 313-326.
- Srinivisan, R., Parthasarathy, S. (1976), *Some Statistical Applications in X-ray Crystallography*, Pergamon Press.
- Stadje, W. (1989), "Exact probability distribution for non-correlated random walk models", *Journal of Statistical Physics*, 56, 415-435.
- Temme, N. M. (1982), "The uniform asymptotic expansion of a class of integrals related to cumulative distribution functions", *SIAM Journal on Mathematical Analysis*, 13, 239-253.
- Wang, S. (1992), "General saddlepoint approximations in the bootstrap", *Statistics & Probability Letters*, 13, 61-66.
- Wang, S. (1993), "Saddlepoint expansions in finite population problems", *Biometrika*, 80, 583-590.
- Wang, S. (1995), "One-step saddlepoint approximations for quantiles", *Computational Statis-*

tics and Data Analysis, 20, 65-74.

Weiss, G. H., Kiefer, J. E. (1983), "The Pearson random walk with unequal step sizes", *Journal of Physics, A: Mathematical and General*, 16, 489-495.